

An Introduction of Transition-Based Dependency Parsing

Hiroki Ouchi

Computational Linguistics Lab, NAIST

2016/04/18 (D-Lec @ Parsing SG)

GOAL

To be able to **explain**

what transition-based dep parsing is

1. Introduction

Dependency Parsing

Input

She kept a cat .
PRP VBD DT NN .



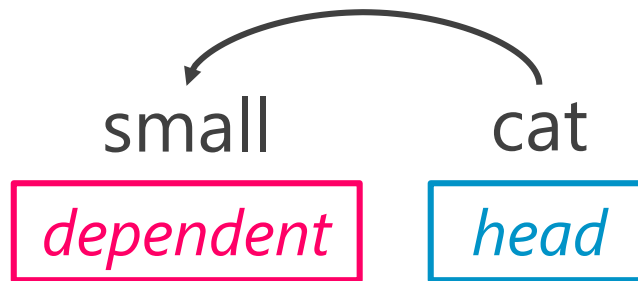
Output

ROOT She kept a cat .

The output shows the words "ROOT She kept a cat ." with dependency arcs. A long arc connects "ROOT" to "kept". A shorter arc connects "kept" to "She". Two arcs connect "kept" to "a" and "kept" to "cat".

Dependency Arcs

Modification Relations

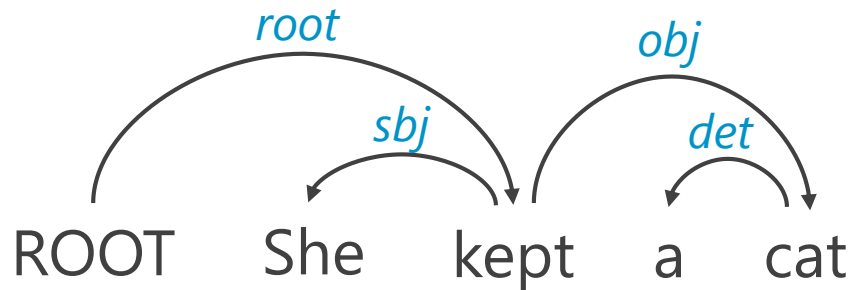


Modified Word = *head*

Modifying Word = *dependent*

Dependency Labels

Detail Syntactic Information



Unlabeled Dep Parsing : Dep arcs

Labeled Dep Parsing : Dep arcs + labels

2. Example

Transition Process

SHIFT

LEFT-ARC

RIGHT-ARC

STACK

ROOT

BUFFER

She kept a cat .

Transition Process

SHIFT

∞

LEFT-ARC

$-\infty$

RIGHT-ARC

$-\infty$

STACK

ROOT

BUFFER

She kept a cat .

Transition Process

SHIFT

∞

LEFT-ARC

$-\infty$

RIGHT-ARC

$-\infty$

STACK

ROOT She

BUFFER

kept a cat .

Transition Process

SHIFT

9.2

LEFT-ARC

0.3

RIGHT-ARC

1.2

STACK

ROOT She

BUFFER

kept a cat .

Transition Process

SHIFT

9.2

LEFT-ARC

0.3

RIGHT-ARC

1.2

STACK

ROOT She kept

BUFFER

a cat .

Transition Process

SHIFT

5.6

LEFT-ARC

6.4

RIGHT-ARC

1.3

STACK

ROOT She kept

BUFFER

a cat .

Transition Process

SHIFT

5.6

LEFT-ARC

6.4

RIGHT-ARC

1.3

STACK

ROOT kept

She



BUFFER

a cat .

Transition Process

SHIFT

8.5

LEFT-ARC

0.4

RIGHT-ARC

0.9

STACK

ROOT kept

She

BUFFER

a cat .

Transition Process

SHIFT

8.5

LEFT-ARC

0.4

RIGHT-ARC

0.9

STACK

ROOT kept a

She



BUFFER

cat .

Transition Process

SHIFT

6.9

LEFT-ARC

2.1

RIGHT-ARC

1.5

STACK

ROOT kept a

She

BUFFER

cat .

Transition Process

SHIFT

6.9

LEFT-ARC

2.1

RIGHT-ARC

1.5

STACK

ROOT kept a cat

She

BUFFER

.

Transition Process

SHIFT

4.6

LEFT-ARC

7.8

RIGHT-ARC

2.6

STACK

ROOT kept a cat

She

BUFFER

.

Transition Process

SHIFT

4.6

LEFT-ARC

7.8

RIGHT-ARC

2.6

STACK

ROOT kept cat

She a

BUFFER

.

Transition Process

SHIFT

3.8

LEFT-ARC

2.2

RIGHT-ARC

8.4

STACK

ROOT kept cat

She a

BUFFER

.

Transition Process

SHIFT

3.8

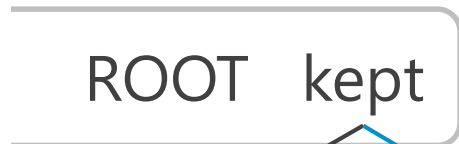
LEFT-ARC

2.2

RIGHT-ARC

8.4

STACK



She cat

a

BUFFER

.

Transition Process

SHIFT

9.1

LEFT-ARC

0.8

RIGHT-ARC

2.7

STACK



She cat

a

BUFFER

.

Transition Process

SHIFT

9.1

LEFT-ARC

0.8

RIGHT-ARC

2.7

STACK

ROOT kept .

She cat

a

BUFFER

Transition Process

SHIFT

$-\infty$

LEFT-ARC

1.8

RIGHT-ARC

8.9

STACK

ROOT kept .

She cat

a

BUFFER

Transition Process

SHIFT

$-\infty$

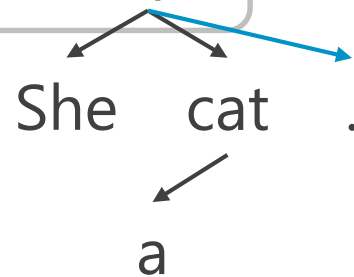
LEFT-ARC

1.8

RIGHT-ARC

8.9

STACK



BUFFER

Transition Process

SHIFT

$-\infty$

LEFT-ARC

0.3

RIGHT-ARC

9.5

STACK

ROOT kept

She cat .
a

```
graph TD; kept[kept] --> She[She]; kept --> cat[cat]; kept --> dot[.]; cat --> a[a];
```

BUFFER

Transition Process

SHIFT

$-\infty$

LEFT-ARC

0.3

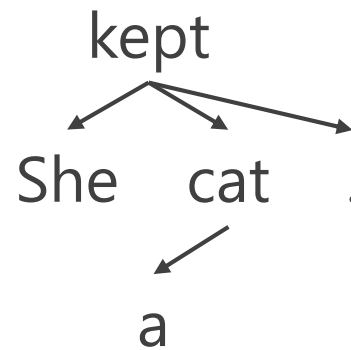
RIGHT-ARC

9.5

STACK



BUFFER



Transition Process

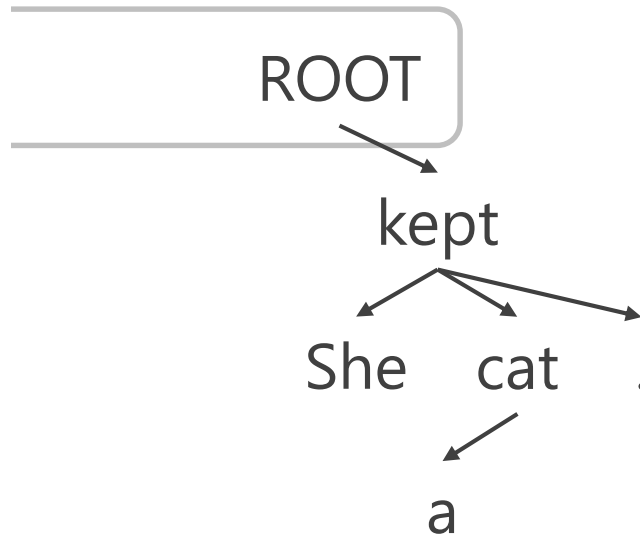
SHIFT

LEFT-ARC

RIGHT-ARC

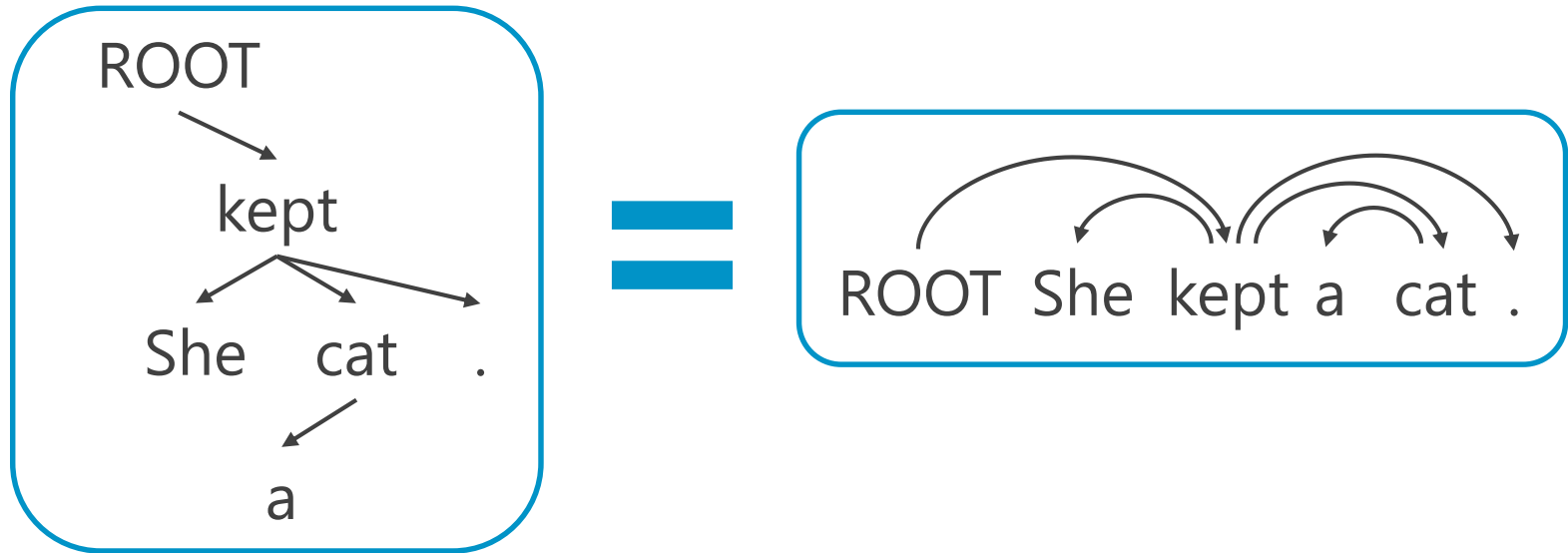
STACK

BUFFER



Terminated

Transition Process



A Resulting Dep Tree

3. Details

Transitions

ROOT She kept

s_1

s_0

a cat .

b_0

Transitions

ROOT She kept

s_1

s_0

a cat .

b_0



SHIFT

$s_0 \leftarrow b_0$

ROOT She kept a

LEFT-ARC

$s_1 \leftarrow s_0$

ROOT kept

She

RIGHT-ARC

$s_1 \rightarrow s_0$

ROOT She

kept

Transitions

ROOT She kept

s_1 s_0

a cat .

b_0

How do we choose next transitions?

SHIFT

$s_0 \leftarrow b_0$

ROOT She kept a

LEFT-ARC

$s_1 \leftarrow s_0$

ROOT kept

She

RIGHT-ARC

$s_1 \rightarrow s_0$

ROOT She

kept

Transition Scores

SHIFT

LEFT-ARC

RIGHT-ARC

5.6

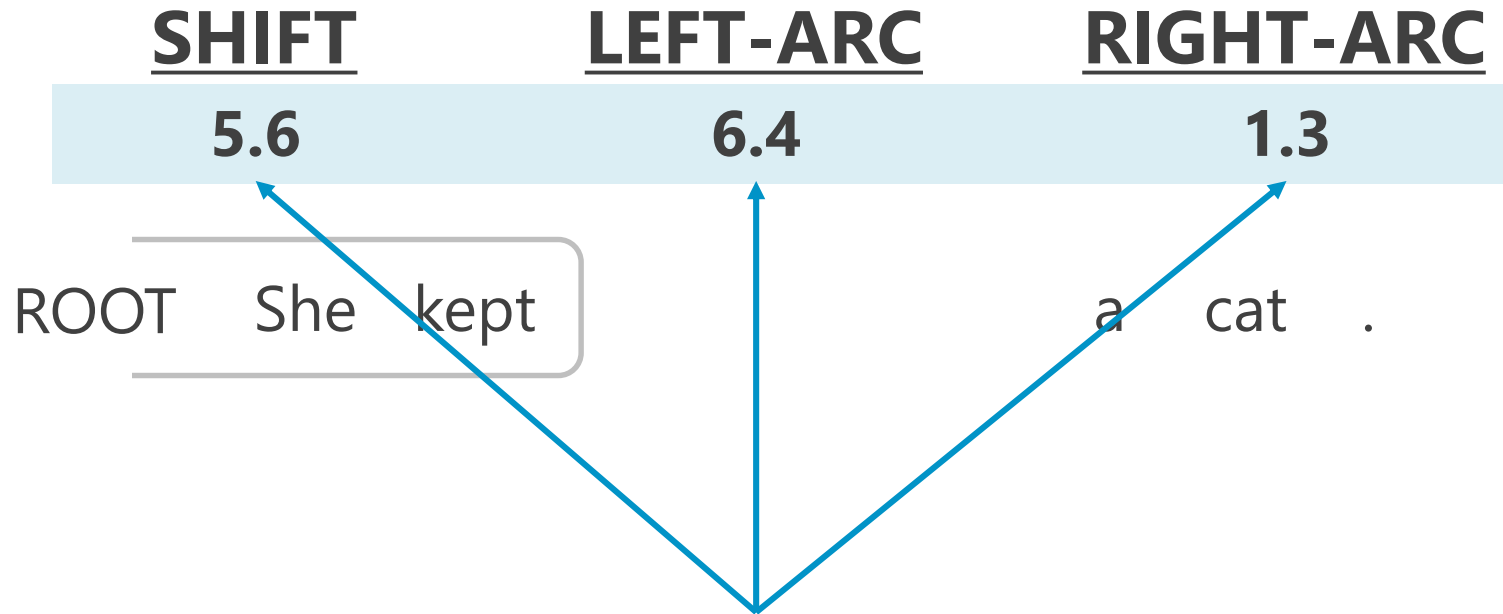
6.4

1.3

ROOT She kept

a cat .

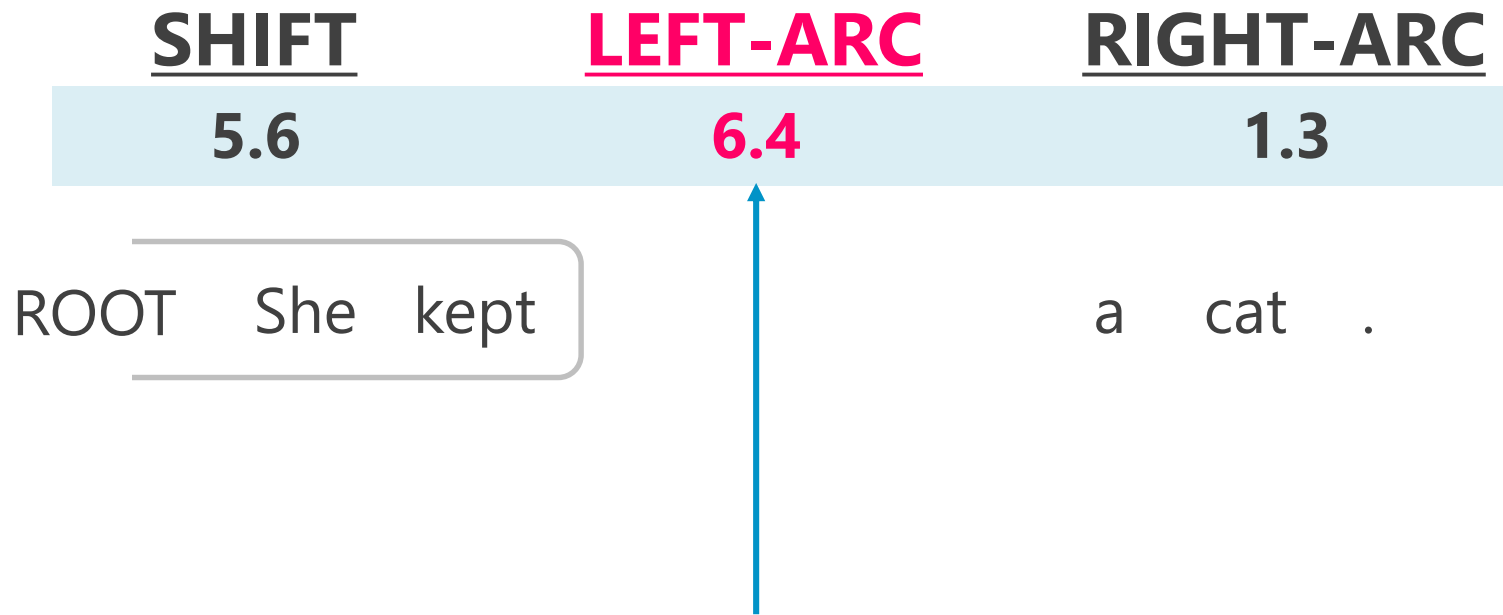
Transition Scores



Machine Learning Techniques

- **Perceptron**
- **Maximum Entropy**
- **Support Vector Machine**
- **Neural Networks**

Transition Scores



Choose the highest scoring transition

4. Summary

Goal : Explanation

Transition-based dep parsing is :

1. Dep trees are predicted with **state transitions**.
2. **States** are represented by the *stack* and *buffer*.
3. **Transitions** are SHIFT, LEFT-ARC, and RIGHT-ARC.
4. As a result of **transition** sequences predicted from **each state**, we can obtain dep trees.

Reference

Tutorials

<http://stp.lingfil.uu.se/~nivre/eacl14.html>

<http://stp.lingfil.uu.se/~nivre/docs/ACLslides.pdf>

Books

